

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_p1ue_p2n model

Run on: April 29, 2004, 08:35:47 ; Search time 0.001 Seconds

(without alignments)
1724.344 Million cell updates/sec

Title: us-09-784-340-2
Perfect score: 2802
Sequence: 1 MRSKDSALVPLQLPFCVCG.....KCFIFSCQKRNKRIKRE 527

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1 segs, 1636 residues
Total number of hits satisfying chosen parameters: 2

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 1 summaries

Command line parameters:
-MODE=frame+ p2n.model -DRV=soft -Q=us-09-784-340-2 -DB=us-10-258-080-11
-SUFFIX=pco -OUT=align2_080-11 -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0 -UNITS=bits
-START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=1 -DOCALIGN=200
-THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=1 -MODE=LOCAL -OUTFMT=spio
-NORM-ext -HAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000 -NCPU=6 -NO_XLTRY
-NEG SCORES=0 -LONGLOG -THREDS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELDP=6 -DELEXT=7

Database : us-10-258-080-11:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	2802	100.0	1636	1	us-10-258-080-11

ALIGNMENTS

RESULT 1
us-10-258-080-11

Alignment Scores:

Pred. No.:	Score:	Percent Similarity:	Best Local Similarity:	Query Match:	Length:	Matches:	Conservative:	Mismatches:	Indels:	Gaps:
0	2802.00	100.00%	100.00%	1	1636	527	0	0	0	0

us-09-784-340-2 (1-527) x us-10-258-080-11 (1-1636)

Qy 1 McTargSerAspLySerAlaLeuValPheLeuLeuGluInLeuPheCysValGlyCys 20
Db 27 ATGAGGCTCGACAGTCAGCTTGGTATTCTGCTCTCGACGCTTCTGTTGGCTGT 86

Qy	21	GlyPheCysGlyValLeuValTrpProCysAspMetSerHisTrpLeuValValys	40
Db	87	GGATTCCTGAGGAAAGTCCTGGTGGCCCTGTGACATGAGCATTGGCTTATATGCAAG	146
Qy	41	ValIleLeuGluGluLeuIleValArgGlyHisGluValThrValLeuThrHisSerLys	60
Db	147	GTCATTCTAGAGAGCTCATATGAGAGAGCCATGAGTACAGTATGACTCATCAAG	206
Qy	61	ProSerLeuIleAspTyrArgLysProSerAlaLeuLysPheGluValValHisMetPro	80
Db	207	CCTTCGTTAATGACTACAGGAGGCTTCGATTTGAATTTGAGGTGCTCATAGCCA	266
Qy	81	GlnAspArgThrGluGluAsnGluIlePheValAspLeuValLeuAsnValLeuProGly	100
Db	267	CAGACACAGACAGAAAGAAATGAAATTTGTTGACTGACTGTAATGCTTCCAGGCG	326
Qy	101	LeuSerThrTrpGlnSerValIleLysLeuAsnAspPhePheValGluIleArgGlyThr	120
Db	327	TTATCACTCGACATCACTTATTAATTAATTAATTAATTAATTAATTAATTAATTAAT	386
Qy	121	LeuLysMetCysGluSerPheIleTyrAsnGlnThrLeuMetLysLysLeuGlnGlu	140
Db	387	TTAAAAATGATGTGAGAGGCTTATCTACATCAGACCTTATGAGAGACTACAGAA	446
Qy	141	ThrAsnTyrAspValMetLeuIleAspProValIleProCysGlyAspLeuMetAlaGlu	160
Db	447	ACCACTACAGATGATATGCTTATAGACCTTGATTCCTGAGAGACTGATGGCTGAG	506
Qy	161	LeuLeuAlaValProPheValLeuThrLeuArgIleSerValGlyLysAsnMetGluArg	180
Db	507	TTGCTGACAGTCCTTTGTGCTCACACTTGAATTTCTGATGAGAGCAATATGAGGCA	566
Qy	181	SerCysGlyLysLeuProAlaProLeuSerTyrValProValProMetThrGlyLeuThr	200
Db	567	AGCTGTGGAAACTTCCAGCTCCACTTCCATATGACTGCTGCTGATGACAGACTAACA	626
Qy	201	AspArgMetThrPheLeuGluArgValLysAsnSerMetLeuSerValLeuPheHisPhe	220
Db	627	GACAGATGACCTTCTGGAAGAGTAAATTCATATGCTTTCAGTTTGTTCACCTTC	686
Qy	221	TrpIleGlnAspTyrAspTyrHisPheTrpGluGluPheTyrSerLysAlaLeuGlyArg	240
Db	687	TGATTCAGGATTCAGACTATCATTTTGGAGAGGTTTATATGATGACATATGAGAG	746
Qy	241	ProThrThrLeuCysGluThrValGlyLysAlaGluIleTrpLeuIleArgThrTyrTrp	260
Db	747	CCCACTACATATATGAGACTGTGAGGAAAGCTGAGATATGCTATATGACATATTTG	806
Qy	261	AspPheGluPheProGlnProTyrGlnProAsnPheGluPheValGlyLeuHisCys	280
Db	807	GATTTGAAATTTCTCAACCATACCAACTTAACCTTATGATTTGTTGAGAGATTCAC	866
Qy	281	LysProAlaLysAlaLeuProLysGluMetGluAsnPheValGlnSerSerGlyLysP	300
Db	867	AAACCTGCCAAAGCTTTGCCAAGAAATGAAATTTTGTCCAGAGTTCCAGGGAGAA	926
Qy	301	GlyIleValValPheSerLeuGlySerLeuPheGluAsnValThrGluGluValAsn	320
Db	927	GGATTTGATGATTTCTCTGGGGCTACTGTTTCAAAATGTTACAGAAAGAAAGCTAT	986
Qy	321	IleIleAlaSerAlaLeuAlaGlnIleProGlnLysValLeuTrpArgTyrLysGlyLys	340
Db	987	ATCATTCCTTCAGGCTTGGCCCAATCCACAGAGAGGTTATGAGAGTTACAAAGAAA	1046
Qy	341	LysProSerThrLeuGlyAlaAsnThrArgLeuTyrAspTrpIleProGlnAsnAspLeu	360
Db	1047	AAACCATCCACATTAAGAGCAATATCTGGCTGATGATGATGATGATGATGATGAT	1106
Qy	361	LeuGlyHisProLysThrLysAlaPheIleThrHisGlyGlyMetAsnGlyIleThrGlu	380
Db	1107	CTTGCTATCCCAAAACCAAGCTTTTATCATCATGATGAGAAATGAGATTCATGAA	1166

```
QY 381 AAlaIleTyrHisGlyValProMetValGlyValProIlePheGlyAspGlnLeuAspAsn 400
Db 1167 GCTATTTCACATGGGGTCCCTATGGGAGTCCCATATTGTGGATCAAGCTTGATAC 1226
QY 401 IleAlaHisMetLysAlaLysGlyValAlaValGlnIleAsnPheLysThrMetThrSer 420
Db 1227 ATAGCTCAGCATGAAAGGCCAAGAGAGACAGCTGTAGAAATTAACCTCAAAACTATGACAAAGC 1286
QY 421 GluAspLeuLeuAlaGlyAlaLeuArgThrValIleThrAspSerSerTyrLysGluAsnAla 440
Db 1287 GAAGATTACTGAGGGGCTTTGAGAACAGTCAATTAACCAATCCCTTATATAAGAGAAATGCT 1346
QY 441 MetArgLeuSerArgIleHisAspGlnProValLysProLeuAspArgAlaValPhe 460
Db 1347 ATGAGATTATCAAGAAATTCACCATGATCAACCTGTAAAGCCCTAGATCGAGCAGTCTTC 1406
QY 461 TrpIleGluPheValMetArgHisLysGlyValAlaLysHisLeuArgSerAlaHisAsp 480
Db 1407 TGGATCGAGTTTGTCAATCCGCCCAAGAGAGCCAAAGCACTGCGATCAAGCTGCCCATGAC 1466
QY 481 LeuThrTrpPheGlnHisTyrSerIleAspValIleGlyPheLeuLeuThrCysValAla 500
Db 1467 CTCACCTGGTTCACGACTACTCTATAGATGTGATGGGTTCTGCTGACCTGTGTGSCA 1526
QY 501 ThrAlaIlePheLeuPheThrLysCysPheLeuPheSerCysGlnLysPheAsnLysThr 520
Db 1527 ACTGCTATATCTCTGTTCACAAATGTTTTTATTTCTGTCAAAAATTTAATTAATAACT 1586
QY 521 ArgLysIleGluLysArgGlu 527
Db 1587 AGAAAGATGAAAAAGAGGAAA 1607
```

Search completed: April 29, 2004, 08:35:49
Job time : 2 secs


```
Qy 381 AlaIleTyrHisGlyValProMetValGlyValProIlePheGlnIleuAspAsn 400
    |||
Db 1167 GCTATTTCATCATGGGGTCCCTATAGTGGAGTTCCCATATTGGATTCAGCTGATAC 1226
    |||
Qy 401 IleAlaHisMetLysAlaLysGlyValAlaValGluIleAsnPhelysThrMetThrSer 420
    |||
Db 1227 ATAGCTCATGAGGCAAGGCAAGGAGCAGCTGTAGAAATAAATTCAAAACTATGACAAAGC 1286
    |||
Qy 421 GluAspLeuLeuArgAlaLeuArgThrValIleThrAspSerSerTyrLysGluAsnAla 440
    |||
Db 1287 GAAGATTACTGAGGGCTTGAGAACAGTCATTACCGATTCCCTTATTAAGAGAAATGCT 1346
    |||
Qy 441 MetArgLeuSerArgIleHisAspGlnProValLysProLeuAspArgAlaValPhe 460
    |||
Db 1347 ATGAGATTATCAAGAAATTCACCATGATCACTGTAAAGCCCTAGATCGAGCAGTCTTC 1406
    |||
Qy 461 TrpIleGluPheValMetArgHisLysGlyValAlaLysHisLeuArgSerAlaAlaHisAsp 480
    |||
Db 1407 TGGATCGAGTTGTGTCATGCCGCCCAAGAGCCCAAGCAGCTGCATCAGTCCCATGAC 1466
    |||
Qy 481 LeuThrTrpPheGlnHisTyrSerIleAspValIleGlyPheLeuLeuThrCysValAla 500
    |||
Db 1467 CTCACCTGGTTCAGCACTACTATATGATGTGATTGGGTTCCTGCTGACCTGTGGCA 1526
    |||
Qy 501 ThrAlaIlePheLeuPheThrLysCysPheLeuPheSerCysGlnLysPheAsnLysThr 520
    |||
Db 1527 ACTGCTATATTCCTGTTCACAAAATGTTTTTATTTTCCTGTCAAAAATTTAATTAAC 1586
    |||
Qy 521 ArgLysIleGluLysArgGlu 527
    |||
Db 1587 AGAAAGTAGAAAAGAGGGA 1607
    |||
```

Search completed: April 29, 2004, 08:33:39
Job time : 3 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using bw model

Run on: April 29, 2004, 08:27:53 ; Search time 4 Seconds
(without alignments)
4.103 Million cell updates/sec

Title: us-09-784-340-1

Perfect score: 2759

Sequence: 1 caaccatcgagatcagtggtg.....ctgtcagcgcttaccagcg 2759

Scoring table: IDENTITY_NUC -1

Gapop 10.0, Gapept 0.5

Searched: 1 seqs, 2974 residues

Total number of hits satisfying chosen parameters: 2

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database: us-10-206-915-521.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	2335	84.6	2974	1	us-10-206-915-521

ALIGNMENTS

RESULT 1
us-10-206-915-521

Query Match 84.6%; Score 2335; DB 1; Length 2974;
Best Local Similarity 96.3%; Pred. No. 0;
Matches 2441; Conservative 0; Mismatches 85; Indels 10; Gaps 5;

QY	12	GATCAGTGTGTGAGGAACTGCATCATGAGTCTGACAAAGCTTGTGATTTCTGC	71
DB	1	GATCAGTGTGTGAGGAACTGCATCATGAGTCTGACAAAGCTTGTGATTTCTGC	60
QY	72	TCCTGACGCTCTTCTGTGTGCTGTGATTTCTGTGGAAGTCTGTGTGCTGTG	131
DB	61	TCCTGACGCTCTTCTGTGTGCTGTGATTTCTGTGGAAGTCTGTGTGCTGTG	120
QY	132	ACATGAGGCAATGGCTTAATGTCAGAGTCAATCTTAAAGAGCTCATTAGAGGCGCAT	191
DB	121	ACATGAGGCAATGGCTTAATGTCAGAGTCAATCTTAAAGAGCTCATTAGAGGCGCAT	180
QY	192	AGGTAAAGTATTTGATCTCACTCAAGCTTCTGTTAATTAATGATCAAGAGGCTTCTGAT	251
DB	181	AGGTAAAGTATTTGATCTCACTCAAGCTTCTGTTAATTAATGATCAAGAGGCTTCTGAT	240
QY	252	TGAATTTGAGGTGTCTCATATGCCACAGAGCAAGAACAGAAATGAATATTTGTTG	311
DB	241	TGAATTTGAGGTGTCTCATATGCCACAGAGCAAGAACAGAAATGAATATTTGTTG	300

QY	312	ACCTAGGCTGTGATGCTTGTGCAAGCTTATGCACTGCGCATCACTTAAATTAATG	371
DB	301	ACCTAGGCTGTGATGCTTGTGCAAGCTTATGCACTGCGCATCACTTAAATTAATG	360
QY	372	ATTTTGTGTGTAATAGAGAACTTTAAATGATGTGTGAGGCTTTATCTCAATC	431
DB	361	ATTTTGTGTGTAATAGAGAACTTTAAATGATGTGTGAGGCTTTATCTCAATC	420
QY	432	AGACGCTTATGAAAGAGCTACAGAGAACTCACTAGATGATGCTTATAGACCTGTGA	491
DB	421	AGACGCTTATGAAAGAGCTACAGAGAACTCACTAGATGATGCTTATAGACCTGTGA	480
QY	492	TTCCCTGTGAGAGCTGATGCTGATGCTTGTGCAATGCTTGTGTGCTCACTTGA	551
DB	481	TTCCCTGTGAGAGCTGATGCTGATGCTTGTGCAATGCTTGTGTGCTCACTTGA	540
QY	552	TTTCTGTAGAGGCAATATGAGGGAAGCTGTGGAAGCTTCCAGCTTCACTTCTATG	611
DB	541	TTTCTGTAGAGGCAATATGAGGGAAGCTGTGGAAGCTTCCAGCTTCACTTCTATG	600
QY	612	TACCTGTGCTTATGACAGACTAAACAGACAGAAATGACCTTCTGGAAGATTAATTT	671
DB	601	TACCTGTGCTTATGACAGACTAAACAGACAGAAATGACCTTCTGGAAGATTAATTT	660
QY	672	CAATGCTTTGAGTTTGTTCATCTTGTGATTCAGATTAAGACTATCAATTTTGGAG	731
DB	661	CAATGCTTTGAGTTTGTTCATCTTGTGATTCAGATTAAGACTATCAATTTTGGAG	720
QY	732	AGTTTATATGAGGCAATGAGAGGCGCACTACATTAATGACATGTGTGGAAGCTG	791
DB	721	AGTTTATATGAGGCAATGAGAGGCGCACTACATTAATGACATGTGTGGAAGCTG	780
QY	792	AGATATGCTTAATAGCAATATTTGGATTTGATTTCCCTCAACATCACTCACTACT	851
DB	781	AGATATGCTTAATAGCAATATTTGGATTTGATTTCCCTCAACATCACTCACTACT	840
QY	852	TTGAGTTTGTGAGAGATTCACCTGTAACCTGCAAGCTTGTCTTAAGAAATGAAA	911
DB	841	TTGAGTTTGTGAGAGATTCACCTGTAACCTGCAAGCTTGTCTTAAGAAATGAAA	900
QY	912	ATTTTGTGCAATTCAGAGGGAAGATGTATTTGTGTGTTTCTGTGGGCTCACTGTTT	971
DB	901	ATTTTGTGCAATTCAGAGGGAAGATGTATTTGTGTGTTTCTGTGGGCTCACTGTTT	960
QY	972	AAATGTACAGAAAGAGCTAATATCATGCTTCAAGCTTGTGCGCAATCCACAGA	1031
DB	961	AAATGTACAGAAAGAGCTAATATCATGCTTCAAGCTTGTGCGCAATCCACAGA	1020
QY	1032	AGGTGTATGAGAGTACAAAGGAAAAAACATCCACATTAAGAGCAATCTGGCTGT	1091
DB	1021	AGGTGTATGAGAGTACAAAGGAAAAAACATCCACATTAAGAGCAATCTGGCTGT	1080
QY	1092	ATGATTTGATACCCCGAATGATCTTGTGTGATCCCAAAAGCTTTATCACTC	1151
DB	1081	ATGATTTGATACCCCGAATGATCTTGTGTGATCCCAAAAGCTTTATCACTC	1140
QY	1152	ATGATGAAATGATGAGATCTAATGAGCTAATTAACAATGGGCTCCCTATAGTGGAGTTC	1211
DB	1141	ATGATGAAATGATGAGATCTAATGAGCTAATTAACAATGGGCTCCCTATAGTGGAGTTC	1200
QY	1212	CCATATTTGTGATGATGCTTGAATACATAGCTCACAATGAGGCAAGAGGAGCTGTAG	1271
DB	1201	CCATATTTGTGATGATGCTTGAATACATAGCTCACAATGAGGCAAGAGGAGCTGTAG	1260
QY	1272	AAATTAATCTCAAACTAATGACAGGGAAGTTTACTGAGGCTTTGAGAACTGATTA	1331
DB	1261	AAATTAATCTCAAACTAATGACAGGGAAGTTTACTGAGGCTTTGAGAACTGATTA	1320
QY	1332	CCGATTTCTCTATTAAGAGATGCTATGAGATTAATCAAGAAATTCACATGATCACTG	1391
DB	1321	CCGATTTCTCTATTAAGAGATGCTATGAGATTAATCAAGAAATTCACATGATCACTG	1380
QY	1392	TAAAGCCCTATGATGAGAGTCTTCTGTGATGAGTTTGTATGCGCCACAAAGAGCA	1451

```
Db 1381 TAAAGCCCTAGATCGACAGATCTTCGATCGAGTTTGTCTACGCGCCACAAAGAGCC 1440
Qy 1452 AGCACTGGGATAGCTGCCCATGACTCTACCTGGTTCAGACACTCTATAGATGTA 1511
Db 1441 AGCACTGGGATAGCTGCCCATGACTCTACCTGGTTCAGACACTCTATAGATGTA 1500
Qy 1512 TTGGGTTCCGCGACCTGGTGGCACTGGTATCTCTGTTTCACAAAATGTTTTTAT 1571
Db 1501 TTGGGTTCCGCGACCTGGTGGCACTGGTATCTCTGTTTCACAAAATGTTTTTAT 1560
Qy 1572 TTTCGTGCAAAATTTTAAACTGAGAAGTAGAAAAGGGAAATAGATCTTCCAA 1631
Db 1561 TTTCGTGCAAAATTTTAAACTGAGAAGTAGAAAAGGGAAATAGATCTTCCAA 1620
Qy 1632 ATTCAAGAAAGACTGATGGGGTATCTGTATATTCAGCCAATAGAAATTTGTGAAA 1691
Db 1621 ATTCAAGAAAGACTGATGGGGTATCTGTATATTCAGCCAATAGAAATTTGTGAAA 1680
Qy 1692 ACCTGTCTATTTTCAATATCTATTTCTGTATTTTATCTTACGTATATAGCTAGATT 1751
Db 1681 ACCTGTCTATTTTCAATATCTATTTCTGTATTTTATCTTACGTATATAGCTAGATT 1740
Qy 1752 CCATGATCAGAGGTGGTGGATATCTCATCTTCGTTGCACTTTCCTAGGGTCTT 1811
Db 1741 CCATGATCAGAGGTGGTGGATATCTCATCTTCGTTGCACTTTCCTAGGGTCTT 1800
Qy 1812 ACTCTCTCTCTCACTTTGTGACACAGACATGAATACATCTAAATTTTCTATTTCTG 1871
Db 1801 ACTCTCTCTCTCACTTTGTGACACAGACATGAATACATCTAAATTTTCTATTTCTG 1860
Qy 1872 ATATCATCTGTTTCATGACGTCAATTACTTCTTAACCTTAAGTATAGGGTGAACCTGCAA 1931
Db 1861 ATATCATCTGTTTCATGACGTCAATTACTTCTTAACCTTAAGTATAGGGTGAACCTGCAA 1920
Qy 1932 TATGCGATTTCCGCGGTTTGGACAAACATGATGTAAAGATGTAAGTAAATATGTAAAT 1991
Db 1921 TATGATTTATTCGCGGTTGCGCCAAACATGATGTAAAGATGTAAGTAAATATGTAAAT 1980
Qy 1992 TCACAAATTCAGTAAACACACAAATCAATGAAAGATCTATGACATTAAGTGTATG 2051
Db 1981 TCACAAATTCAGTAAACACACAAATCAATGAAAGATCTATGAGATTAAGTGTATG 2040
Qy 2052 AGTAACATATGATTTTCTTTTCAATTAATTAAGCCCTTCTACATACCCAGCATTAC 2111
Db 2041 AGAAACATATGATTTTCTTTTCAATTAATTAAGCCCTTCTACATACCCAGCATTAC 2100
Qy 2112 TGATCTCAGACATGAATGCTTAAATGACATAGGGCATTAACCTCAGAAATAGTTGC 2171
Db 2101 TGATCTCAGAAATTAATGCTTAAATGACATAGGCATTAATGCTTGAAGAAATTTGC 2160
Qy 2172 TATATTTCCACATACCTCATCTAGATGTACATAGCTTCTGCAATCACTTAATGTA 2231
Db 2161 TGTATTTCCATAGACTCATCTAGATGTACATAGCTTCTGCAATCACTTAATGTA 2220
Qy 2232 CA-TTTTTGTGTGTTCTGATGATAATAGACAGTTCTTATTAATGTCCTCAATATA 2290
Db 2221 TACTTTTTCTGTTTCTGATGATAATAGACAGTTCTTCAATGATGCAATATACA 2280
Qy 2291 AAAGAAACT-GAAATTTTCTTCACTAGAGAAAATGTCATTAAGATTTCAAGTTAAACAG 2349
Db 2281 AAAGAAACTAATTTTTTTCTCATAAGAGAACATGTCAGTAAGATTTCAAGGTGAACAG 2340
Qy 2350 ATTATTTTGAATTAAGTAAACATTAAGAAATATGTATGTTTCTGATTTTATATAAAT 2409
Db 2341 AATATTTTGGGATTAAGTAAATTAATGATGTTGATTAATTAAGTTTATATAA 2399
Qy 2410 TTTAATGATGATCACTT-----GATTAATGTCATTTCTT-AAATGATGATATAC 2462
Db 2400 TTTAATGATGATCACTTAAAGAGATTTATATGTTTATCTTTAAATGATGATATAC 2459
Qy 2463 TCATATTTCTTATCTCTATTAATCAAAAGTAAATTTACTGTAGAAAAATTAAGATGCT 2522
```

```
Db 2460 TCATATTTCTTATCTCTATTAATCAAAAGTAAATTTACTGTAGAAAAATTAAGATGCT 2519
Qy 2523 TGTTCGAAAGTAAAA 2538
Db 2520 TGTTCGAAAGTAAAA 2535
```

Search completed: April 29, 2004, 08:27:58
Job time : 5 secs

Db 241 TGAATTGAGGTGGTCCATATGCCACAGACAGAACAGAAATGAATATTGTTG 30

QY	312	ACCTAGCTGATATGCTTGCCAGGCTTATCAACCTGGGCAATCACTAATATAAATAATG	37
Db	301	ACCTAGCTGATATGCTTGCCAGGCTTATCAACCTGGGCAATCACTAATATAAATAATG	36
QY	372	ATTTTGTGTGAATAATAGAGNACTTTAAATAATGATGTGTGAGAGCTTATCTAATC	43
Db	361	ATTTTGTGTGAATAATAGAGNACTTTAAATAATGATGTGTGAGAGCTTATCTAATC	42
QY	432	AGAGCTTATGAAGAGCTACAGAAACCACTAGATGTATGCTTATAGACCTGTGA	49
Db	421	AGAGCTTATGAAGAGCTACAGAAACCACTAGATGTATGCTTATAGACCTGTGA	48
QY	492	TTCCCGTGAACCGTATGAGCTGAGTGTCTTGACATCCCTTTGTGTCTACACTTGA	55
Db	481	TTCCCGTGAACCGTATGAGCTGAGTGTCTTGACATCCCTTTGTGTCTACACTTGA	54
QY	552	TTTCTGTAGAGCAATATGAGCGAAGCTGTGGGAACTTCAGGCTCCACTTCCATG	61
Db	541	TTTCTGTAGAGCAATATGAGCGAAGCTGTGGGAACTTCAGGCTCCACTTCCATG	60
QY	612	TACCTGTGCTATGACAGAGCTAACAGACAGATGACCTTCTGAAAGATATAAAT	67
Db	601	TACCTGTGCTATGACAGAGCTAACAGACAGATGACCTTCTGAAAGATATAAAT	66
QY	672	CAATGCTTCAAGTTTGTGTCCACTTCTGATTCAGATTAACAATATCTTTTGGAG	73
Db	661	CAATGCTTCAAGTTTGTGTCCACTTCTGATTCAGATTAACAATATCTTTTGGAG	72
QY	732	AGTTTATAGTAAAGCAATAGGAGGCGCACTAACATATATAGTACATGTGGGAAAGCTG	79
Db	721	AGTTTATAGTAAAGCAATAGGAGGCGCACTAACATATATAGTACATGTGGGAAAGCTG	78
QY	792	AGATATGCTAATACGAATATGGAATTTTGAATTTCTGACACATACCAACTAAT	85
Db	781	AGATATGCTAATACGAATATGGAATTTTGAATTTCTGACACATACCAACTAAT	84
QY	852	TTGATTTGTGTGAGAGATTGCACTGTAACTGCGCAAGCTTGTGCTTAAGAAATGAAA	91
Db	841	TTGATTTGTGTGAGAGATTGCACTGTAACTGCGCAAGCTTGTGCTTAAGAAATGAAA	90
QY	912	ATTTTGTCCAGAGTTCAAGGGAGATGTGATTTGTGTGTCTCTGAGGCTCACTGTTC	97
Db	901	ATTTTGTCCAGAGTTCAAGGGAGATGTGATTTGTGTGTCTCTGAGGCTCACTGTTC	96
QY	972	AAAAATGTACAGAGAAAGGCTAATATCATGTGCTTCAGGCTTGCAGATCCACAGA	103
Db	961	AAAAATGTACAGAGAAAGGCTAATATCATGTGCTTCAGGCTTGCAGATCCACAGA	102
QY	1032	AGGTGTATGAGAGTACAAAGAAAAAACATCCACTTATGAGCCATATCTGGCTGT	109
Db	1021	AGGTGTATGAGAGTACAAAGAAAAAACATCCACTTATGAGCCATATCTGGCTGT	108
QY	1092	ATGATGTGATACCCCAAGATGATCTTTCTGTGTCTATCCCAAAACCAAGCTTATATCTC	115
Db	1081	ATGATGTGATACCCCAAGATGATCTTTCTGTGTCTATCCCAAAACCAAGCTTATATCTC	114
QY	1152	ATGTGTAGATAGATGGATCTATGAGCTATTTACATGGGGGTCCCTATGTTGGAGATTC	121
Db	1141	ATGTGTAGATAGATGGATCTATGAGCTATTTACATGGGGGTCCCTATGTTGGAGATTC	120
QY	1212	CCATATTTGTGATCAGCTTGATTAATATGCTGATCATGAAAGCCAAAGAGCAGCTGATG	127
Db	1201	CCATATTTGTGATCAGCTTGATTAATATGCTGATCATGAAAGCCAAAGAGCAGCTGATG	126
QY	1272	AAATTAACCTCAAAACCTATGACAAAGCAATTTACTGAGGGCTTTGAGACAAGCTATTA	133
Db	1261	AAATTAACCTCAAAACCTATGACAAAGCAATTTACTGAGGGCTTTGAGACAAGCTATTA	132
QY	1332	CCGATTTCTCTTATTAAGAGAAATGCTATGAGATTAATCAAGAAATCAACATATCAACTG	139
Db	1321	CCGATTTCTCTTATTAAGAGAAATGCTATGAGATTAATCAAGAAATCAACATATCAACTG	138
QY	1392	TAAAGCCCTAGATTCAGACAGCTTCTTGATTCAGATTTGTCTATGCGCCAAAGAGAGCA	145

```
DB      1381 TAAAGCCCTAGATGAGAGAGTCTTCGATGAGTTTGTCAAGCCGCAAGAGAGCCA 1440
QY      1452 AGCACTGCGATCAGCTGCCATGACCTCACTGGTTCCAGCACTACTATAGATGTA 1511
DB      1441 AGCACTGCGATCAGCTGCCATGACCTCACTGGTTCCAGCACTACTATAGATGTA 1500
QY      1512 TTGGGTTCTGCTGACCTGNGGCACTGCTAATTCTGTTCAGAAAATGTTTTAT 1571
DB      1501 TTGGGTTCTGCTGACCTGNGGCACTGCTAATTCTGTTCAGAAAATGTTTTAT 1560
QY      1572 TTTCCTGTCAAAAATTTAATAAACTAGAAAGATAGAAAAGGGAATAGATCTTCCAA 1631
DB      1561 TTTCCTGTCAAAAATTTAATAAACTAGAAAGATAGAAAAGGGAATAGATCTTCCAA 1620
QY      1632 ATTCAGAAAAGACCTG 1647
DB      1621 ATTCAGAAAAGACCTG 1636
```

Search completed: April 29, 2004, 08:30:21
Job time : 3 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 29, 2004, 08:22:55 ; Search time 0.001 Seconds
(without alignments)
277.729 Million cell updates/sec

Title: us-09-784-340-2

Perfect score: 2802
Sequence: 1 MRSQKSLVFLQLQFCVGC.....KCFLFSCQKFKTKRIEKE 527

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1 seqs, 527 residues

Total number of hits satisfying chosen parameters: 1

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 1 summaries

Database : us-10-206-915-522:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2802	100.0	527	1	us-10-206-915-522

ALIGNMENTS

RESULT 1
us-10-206-915-522

Query Match 100.0%; Score 2802; DB 1; Length 527;

Best Local Similarity 100.0%; Pred. No. 0;
Matches 527; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	MRSQKSLVFLQLQFCVGC	PCVGVLP	PCDMSHMLNVKYL	IBELIVRGHVTVL	THSK	60
Db	1	MRSQKSLVFLQLQFCVGC	PCVGVLP	PCDMSHMLNVKYL	IBELIVRGHVTVL	THSK	60
Qy	61	PSLIDYRKPSALKEEVV	MPQDRTEEN	EIFVDLALNVLP	GLSTWQSVI	KINDFVEIRGT	120
Db	61	PSLIDYRKPSALKEEVV	MPQDRTEEN	EIFVDLALNVLP	GLSTWQSVI	KINDFVEIRGT	120
Qy	61	PSLIDYRKPSALKEEVV	MPQDRTEEN	EIFVDLALNVLP	GLSTWQSVI	KINDFVEIRGT	120
Db	61	PSLIDYRKPSALKEEVV	MPQDRTEEN	EIFVDLALNVLP	GLSTWQSVI	KINDFVEIRGT	120
Qy	121	LKMWCESTINQTLMK	QLENTYDVM	LDPYIPCGDLMAEL	LAIPVTLT	ISVGNNR	180
Db	121	LKMWCESTINQTLMK	QLENTYDVM	LDPYIPCGDLMAEL	LAIPVTLT	ISVGNNR	180
Qy	121	LKMWCESTINQTLMK	QLENTYDVM	LDPYIPCGDLMAEL	LAIPVTLT	ISVGNNR	180
Db	121	LKMWCESTINQTLMK	QLENTYDVM	LDPYIPCGDLMAEL	LAIPVTLT	ISVGNNR	180
Qy	181	SCGLPAPLSYVPVMT	GLTDRMT	FLERVKNSML	SVLFHFWI	ODYDHFMEEFYSKALGR	240
Db	181	SCGLPAPLSYVPVMT	GLTDRMT	FLERVKNSML	SVLFHFWI	ODYDHFMEEFYSKALGR	240
Qy	181	SCGLPAPLSYVPVMT	GLTDRMT	FLERVKNSML	SVLFHFWI	ODYDHFMEEFYSKALGR	240
Db	181	SCGLPAPLSYVPVMT	GLTDRMT	FLERVKNSML	SVLFHFWI	ODYDHFMEEFYSKALGR	240
Qy	241	PTTLCETVQKAEI	WLIRTYMD	FEFPQYPQ	PNFEFVG	ILHCKPAKALPKEMENFVQSSGD	300
Db	241	PTTLCETVQKAEI	WLIRTYMD	FEFPQYPQ	PNFEFVG	ILHCKPAKALPKEMENFVQSSGD	300
Qy	241	PTTLCETVQKAEI	WLIRTYMD	FEFPQYPQ	PNFEFVG	ILHCKPAKALPKEMENFVQSSGD	300
Db	241	PTTLCETVQKAEI	WLIRTYMD	FEFPQYPQ	PNFEFVG	ILHCKPAKALPKEMENFVQSSGD	300

Qy	301	GIVFSLGSLFQNTTE	EKANIIASALA	QIPQKVL	MRYGKGGK	PSLIGANTRL	YDWI	POND	360
Db	301	GIVFSLGSLFQNTTE	EKANIIASALA	QIPQKVL	MRYGKGGK	PSLIGANTRL	YDWI	POND	360
Qy	361	LGHKTKAPITTHG	MNGIYEAI	YHGVMP	GVPIEGD	LDNI	IAHMKAGAA	VEINFKTMTS	420
Db	361	LGHKTKAPITTHG	MNGIYEAI	YHGVMP	GVPIEGD	LDNI	IAHMKAGAA	VEINFKTMTS	420
Qy	421	EDLIRALRTVIT	DSSSYKENA	MRSLRH	DDQPVKPL	DRAVFI	EFVMMHKGAK	HLRSAHD	480
Db	421	EDLIRALRTVIT	DSSSYKENA	MRSLRH	DDQPVKPL	DRAVFI	EFVMMHKGAK	HLRSAHD	480
Qy	481	LTFPQHSIDVIG	FLTCVATA	IFLFTK	CFLFSCQK	FNKTRKIEKE	527		
Db	481	LTFPQHSIDVIG	FLTCVATA	IFLFTK	CFLFSCQK	FNKTRKIEKE	527		

Search completed: April 29, 2004, 08:22:56
Job time : 1 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 29, 2004, 08:25:24 ; Search time 0.001 Seconds
(without alignments)
277.729 Million cell updates/sec

Title: us-09-784-340-2

Perfect score: 2802

Sequence: 1 MRSDKSLVLLQLFCVGC.....KCFLFSCQKFNKTRKIEKRE 527

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1 seqs, 527 residues

Total number of hits satisfying chosen parameters: 1

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 1 summaries

Database: us-10-258-080-1:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	2802	100.0	527	1	us-10-258-080-1

ALIGNMENTS

RESULT 1
us-10-258-080-1

Query Match 100.0%; Score 2802; DB 1; Length 527;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 527; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MRSDKSLVLLQLFCVGC	CGKVLWPCDSHMLNVKYLIELIVRGHEVTLTHSK	60
DB	1	MRSDKSLVLLQLFCVGC	CGKVLWPCDSHMLNVKYLIELIVRGHEVTLTHSK	60
QY	61	PSLIDYRKPSALKFEVVM	PODRTEENEIFVDLALNVLPGLSTWQSVIKLNDFFVEIRGT	120
DB	61	PSLIDYRKPSALKFEVVM	PODRTEENEIFVDLALNVLPGLSTWQSVIKLNDFFVEIRGT	120
QY	121	LKMMCBSEFIYNOTLMK	LOETNYDVMIDPVIKPGDLMAELIAPFVLTIRISVGMMER	180
DB	121	LKMMCBSEFIYNOTLMK	LOETNYDVMIDPVIKPGDLMAELIAPFVLTIRISVGMMER	180
QY	181	SCGKLPAPLSYVPVMT	GLTDRMTFLERVKNSMLSVLFHFMIQDYDHFMEEFYSKALGR	240
DB	181	SCGKLPAPLSYVPVMT	GLTDRMTFLERVKNSMLSVLFHFMIQDYDHFMEEFYSKALGR	240
QY	241	PTTLCETVGAELIWL	IRTYWDFEFPQYPQNFEEVGLHCKPAKALPKEMENFVQSSGED	300
DB	241	PTTLCETVGAELIWL	IRTYWDFEFPQYPQNFEEVGLHCKPAKALPKEMENFVQSSGED	300

QY	301	GIWFSIGSLFQNVTEEK	NIASALAQIPQKYLTRYKGGKPPSTLGANTRLYDMPOND	360
DB	301	GIWFSIGSLFQNVTEEK	NIASALAQIPQKYLTRYKGGKPPSTLGANTRLYDMPOND	360
QY	361	LGHPKTKAFITTHGMM	GIYEAIVHGVPMGVPIFGDQDNIAMKAKGAABEINFKTMTS	420
DB	361	LGHPKTKAFITTHGMM	GIYEAIVHGVPMGVPIFGDQDNIAMKAKGAABEINFKTMTS	420
QY	421	EDLLRALRTVITDSS	SYKENMRLSRJHHDQPVKPLDRAVFWIEFWMHKAKHLRSAHD	480
DB	421	EDLLRALRTVITDSS	SYKENMRLSRJHHDQPVKPLDRAVFWIEFWMHKAKHLRSAHD	480
QY	481	LTFQOHSIDVIGFLTC	VAATAIFLFTKCPFPSCQKFNKTRKIEKRE	527
DB	481	LTFQOHSIDVIGFLTC	VAATAIFLFTKCPFPSCQKFNKTRKIEKRE	527

Search completed: April 29, 2004, 08:25:25
Job time: 1 sec

Db 353 DMTFPODLGHPKTAFTTHGNGTNGIYEALYHGVPMVGPMPADDPNDIAHMKAGAAVE 412
Qy 413 INKFTWSEDLRALRTVITDSSYKENAMRLSRHHDDPVKPLDRAVFWIEFVRHKGAK 472
Db 413 VNNTMTSVLDLALRTVINEPSEYKENAMRLSRHHDDPVKPLDRAVFWIEFVRHKGAK 472
Qy 473 HURSAADLTWFOHYSIDVIGFLLTCVATFIFLTKCFLESCOKFNKTRIEKRE 527
Db 473 HURSAADLTWFOHYSIDVIGFLLTCVATFIFLTKCFLESCOKFNKTRIEKRE 527

RESULT 4
075614 PRELIMINARY; PRT: 528 AA.

AC 075614;
DT 01-NOV-1998 (TREMBLrel. 08, Created)
DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)
DE UDP GLUCURONOSYLTRANSFERASE 2B4 PRECURSOR.
GN UGT2B4.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93326164; PubMed=8333863;
RA Jin C.J., Miners J.O., Lillywhite K.J., Mackenzie P.I.;
RT "cDNA cloning and expression of two new members of the human liver
UDP-glucuronosyltransferase 2B subfamily."
RL Biochem. Biophys. Res. Commun. 194:496-503(1993).
RN [2]
RP SEQUENCE FROM N.A.
RA Mackenzie P.I.;
RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF081793; AAC32272.1;
DR InterPro: IPR002213;
DR Pfam: PF00201; UDPGT_1.
DR PROSITE: PS00375; UDPGT_1.
KW transferase.
SQ SEQUENCE 528 AA: 60497 MW; B6F5BA0705E952C4 CRC64;

Query Match 61.1%; Score 1711; DB 4; Length 528;
Best Local Similarity 62.1%; Pred. No. 4e-134;
Matches 328; Conservative 64; Mismatches 132; Indels 4; Gaps 3;

Qy 3 SDKSALVFLLLQDFC-VGCGFCGKLVMPCDMSHMLNVKILIEELVGRHEVTVLHSHK 61
Db 2 SMKWTSAALLTLOLSCYFSSGCGKLVMPCTERSHMANITLIDELVQGHVTVLASSAS 61
Qy 62 SLIDYRKPALKFEVYHMPQDRTEENELFVDALN--VLPGLSTWOSVTKLNDFEVEING 119
Db 62 ISFDNSSTLKFVYPSLTKTEFEDIIKQLVKRMALPKDTPMSYLSQVDEIMMTND 121
Qy 120 TLKMMCEFIYNOTLAKLOETNYVMDLPYIPCGDLMAELLANPVYLTIRISVGNNK 179
Db 122 ILRKCKIVSKKMKLKOESRFVDLADAFPRGELLAEILKIPFYVSLRFSFGYALE 181
Qy 180 RSCGKLPLPLSVPMGTGLDRMTFLERVKNSMLVLFHFIQDYDHFHWEFEFSKALG 239
Db 182 KHSGLLPSPVYPMVSELDQMTFLERVKNMIVLYFEFQFLDKMQDFSEVYL 241
Qy 240 RPTLTCTVGAKEIWLIRTYWDFEFPQYQPNFEFVGLHCKPAKALPKEMENFVSSGE 299
Db 242 RPTLTCTVGAKEIWLIRTYWDFEFPQYQPNFEFVGLHCKPAKALPKEMENFVSSGE 301
Qy 300 DGIYVFSGLFQNVTEKANIISALAOIPQKYLMMKKGKPPSTLGANTRLYMIPOND 359
Db 302 NGVYVFSGLFQNVTEKANIISALAOIPQKYLMMKKGKPPSTLGANTRLYMIPOND 361
Qy 360 LIGHPKTAFTTHGNGTNGIYEALYHGVPMVGPMPADDPNDIAHMKAGAAVEIFKMT 419
Db 360 LIGHPKTAFTTHGNGTNGIYEALYHGVPMVGPMPADDPNDIAHMKAGAAVEIFKMT 419

Db 362 LIGHPKTAFTTHGNGTNGIYEALYHGVPMVGPMPADDPNDIAHMKAGAAVEIFKMT 421
Qy 420 SEDLLALRTVITDSSYKENAMRLSRHHDDPVKPLDRAVFWIEFVRHKGAKHLSAAH 479
Db 422 STDLLALRTVITDSSYKENAMRLSRHHDDPVKPLDRAVFWIEFVRHKGAKHLSAAH 481
Qy 480 DLTWFOHYSIDVIGFLLTCVATFIFLTKCFLESCOKFNKTRIEKRE 527
Db 482 DLTWFOHYSIDVIGFLLTCVATFIFLTKCFLESCOKFNKTRIEKRE 527

RESULT 5
09564 PRELIMINARY; PRT: 527 AA.

AC 09564;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE OLFACTORY UDP GLUCURONOSYLTRANSFERASE.
GN Mus musculus (Mouse).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BALE/C;
RA Heydel J.M.;
RT "Mus musculus mRNA for olfactory UDP glucuronosyltransferase."
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF184901; AAG17003.1;
KW transferase.
SQ SEQUENCE 527 AA: 59878 MW; 136FB2A648F39C75 CRC64;

Query Match 60.9%; Score 1705.5; DB 11; Length 527;
Best Local Similarity 61.9%; Pred. No. 1.1e-133;
Matches 326; Conservative 70; Mismatches 110; Indels 21; Gaps 5;

Qy 13 LQLEVCVCGFCGKLVMPCDMSHMLNVKILIEELVGRHEVTVLHSHKPSAL 72
Db 10 LQISLGMSSGKLVMPCDMSHMLNVKILIEELVGRHEVTVLHSHKPSAL 69
Qy 73 KEVYHMPQDRTEENELFVDALNVL--PGLSTWOSVTKLNDFEVEING 124
Db 70 TFEIYVFPFKEKIESIKQFVLTWLENRPSSTIMTFYKEMAKVIEFHLVSRG---I 125
Qy 125 CSEFIYNOTLAKLOETNYVMDLPYIPCGDLMAELLANPVYLTIRISVGNNK 164
Db 126 CDEVLNEMKLSKLOKEFEVLLSDPVFPGDILVALKLGIPFYVSLRFSFGYALE 165
Qy 185 LPAPLSYVPMGTGLDRMTFLERVKNSMLVLFHFIQDYDHF--WEFYSKALGR 240
Db 186 VPPRPVYVALLSELDDQSPADRKN-----FISYRMDYMEFTLMKQNDSTYTRALGR 240
Qy 241 RPTLTCTVGAKEIWLIRTYWDFEFPQYQPNFEFVGLHCKPAKALPKEMENFVSSGE 300
Db 241 RPTLTCTVGAKEIWLIRTYWDFEFPQYQPNFEFVGLHCKPAKALPKEMENFVSSGE 300
Qy 301 GIVYVFSGLFQNVTEKANIISALAOIPQKYLMMKKGKPPSTLGANTRLYMIPOND 360
Db 301 GIVYVFSGLFQNVTEKANIISALAOIPQKYLMMKKGKPPSTLGANTRLYMIPOND 360
Qy 361 LIGHPKTAFTTHGNGTNGIYEALYHGVPMVGPMPADDPNDIAHMKAGAAVEIFKMT 420
Db 361 LIGHPKTAFTTHGNGTNGIYEALYHGVPMVGPMPADDPNDIAHMKAGAAVEIFKMT 420
Qy 421 EDLLALRTVITDSSYKENAMRLSRHHDDPVKPLDRAVFWIEFVRHKGAKHLSAAH 480
Db 421 EDLLALRTVITDSSYKENAMRLSRHHDDPVKPLDRAVFWIEFVRHKGAKHLSAAH 480
Qy 481 LTMFOHYSIDVIGFLLTCVATFIFLTKCFLESCOKFNKTRIEKRE 527
Db 481 LTMFOHYSIDVIGFLLTCVATFIFLTKCFLESCOKFNKTRIEKRE 527

QY 361 LGHPKRAFTTHGNGNGIYEAITYHGVPMVGPPIFGQDLNDINAMKGAAGAEINFTKMTS 420
DB 361 LGHPKRAFTTHGNGNGIYEAITYHGVPMVGPPIFGQDLNDINAMKGAAGAEINFTKMTS 420
QY 421 EDLLRALRTVITDSSYKENAMRLSRHHDDQVVKPLDRAVFWIEFVNRHKGAKHLRSAAHD 480
DB 421 EDLLRALRTVITDSSYKENAMRLSRHHDDQVVKPLDRAVFWIEFVNRHKGAKHLRSAAHD 480
QY 481 LTMFOHYSIDVIGFLTCVATAIFLTKCFLESCQKFNKTRKIEKRE 527
DB 481 LTMFOHYSIDVIGFLTCVATAIFLTKCFLESCQKFNKTRKIEKRE 527

RESULT 4
A:Accession: A42233
glucuronosyltransferase (EC 2.4.1.17) 2 - rat

C:Species: Rattus norvegicus (Norway rat)
C:Date: 19-Jun-1992 #sequence_revision 18-Sep-1992 #text_change 29-Sep-1999
C:Accession: A42233; A24324
R:MacKenzie, P.I.; Rodbourn, L.
J. Biol. Chem. 265, 11328-11332, 1990
A:Title: Organization of the rat UDP-glucuronosyltransferase, UDRGTR-2, gene and character
A:Reference number: A42233; MUID:90293083
A:Accession: A42233
A:Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-529 <MAC>
A:Cross-references: GB:J05482
A:Note: the authors translated the codon GTA for residue 57 as B, and GTC for residue 11
R:MacKenzie, P.I.
J. Biol. Chem. 261, 6119-6125, 1986
A:Title: Rat liver UDP-glucuronosyltransferase. Sequence and expression of a cDNA encoded
A:Reference number: A24324; MUID:86196018
A:Accession: A24324
A:Molecule type: mRNA
A:Residues: 1-407, 'V', 409-529 <MAC>
A:Cross-references: GB:M3506; NID:9207580; PIDN:AAA42313.1; PID:9207581
A:Experimental source: liver
C:Superfamily: glucuronosyltransferase
C:Keywords: glycosyltransferase; hexosyltransferase; transmembrane protein

Query Match 60.2%; Score 1687; DB 2; Length 529;
Best Local Similarity 60.0%; Pred. No. 1.8e-125;
Matches 316; Conservative 74; Mismatches 133; Indels 4; Gaps 2;

QY 3 SDKALVALLQFC-VCCGFCGKVLVPCDMSHLANKVLEELVIGHEVTVLTHSKP 61
DB 2 SMKOTSVFLQLQICYFRGACGKVLVWPTESHMINIKILNELAQGHVETVLYSSAS 61
QY 62 SLIDYRPSALKFEVHMPDORTENEIF--VDLALNVLPGLSTWGVITKLNDFVEIR 118
DB 62 ILIPTFESSINFEIYVPLSKSDLETSEFAKWIDEMTRDFETLSIMWYYSKQKVFNEYS 121
QY 119 GTLMMCESFIYNOTLKKKLOETNYDMLIDPVIIPCGDIAELLAVPFLTLRTISVGNM 178
DB 122 DVVENLKALIMNKSMLKKLOGSDVILADAVPCGELLALLLKTLPLVYSLRCPGRC 181
QY 179 ERSGCKLPAPLSTVYPVMTGTDRTFLERVKNSMLSVLFHFWIQDYHFWEEFYSKAL 238
DB 182 EKESGGRLPSPSYVVALSELSDRMTFVERKNNLMILYDFWFQPFKEKSMOFSYDVL 241
QY 239 GRPTLLETGVKAEIWLIRTYMDEFPPOPYPNFEFVGLHCKPAKALPKRMENFVSSG 298
DB 242 GRPTLLETGVKAEIWLIRTYMDEFPPOPYPNFEFVGLHCKPAKALPKRMENFVSSG 298
QY 299 EDGIVFSLGSFQNTVEKANIATASALAOIPQKVLWRKYGKPKSTLGANTRLDWTIPON 358
DB 302 EHGIVFSLGSFQNTVEKANIATASALAOIPQKVLWRKYGKPKSTLGANTRLDWTIPON 358
QY 359 DLGHPKRAFTTHGNGNGIYEAITYHGVPMVGPPIFGQDLNDINAMKGAAGAEINFTKMTS 418
DB 362 DLGHPKRAFTTHGNGNGIYEAITYHGVPMVGPPIFGQDLNDINAMKGAAGAEINFTKMTS 421

QY 419 TSEDLLRALRTVITDSSYKENAMRLSRHHDDQVVKPLDRAVFWIEFVNRHKGAKHLRSAA 478
DB 422 STGGLTALKIYVNDPSYKENAMRLSRHHDDQVVKPLDRAVFWIEFVNRHKGAKHLRSAA 481
QY 479 HDLTMFOHYSIDVIGFLTCVATAIFLTKCFLESCQKFNKTRKIEKRE 525
DB 482 HDLTMFOHYSIDVIGFLTCVATAIFLTKCFLESCQKFNKTRKIEKRE 528

RESULT 5
A:Accession: S11309
glucuronosyltransferase (EC 2.4.1.17) - human

N:Alternate names: estriol UDPglucuronosyltransferase
C:Species: Homo sapiens (man)
C:Date: 21-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 29-Sep-1999
C:Accession: S11309
R:Coffman, B.L.; Tephy, T.R.; Irshaid, Y.M.; Green, M.D.; Smith, C.; Jackson, M.R.; Arch. Biochem. Biophys. 281, 170-175, 1990
A:Title: Characterization and primary sequence of a human hepatic microsomal estriol
A:Reference number: S11309; MUID:90343358
A:Accession: S11309
A:Status: Preliminary
A:Molecule type: mRNA
A:Residues: 1-523 <COR>
A:Cross-references: GB:U06641; NID:9458398; PIDN:AAA83406.1; PID:9458399
A:Note: the authors translated the codon AGG for residue 412 as Thr and ACC for resi.
C:Superfamily: glucuronosyltransferase
C:Keywords: glycosyltransferase; hexosyltransferase; transmembrane protein

Query Match 60.1%; Score 1684; DB 2; Length 523;
Best Local Similarity 61.7%; Pred. No. 3e-125;
Matches 325; Conservative 72; Mismatches 118; Indels 12; Gaps 5;

QY 9 VELLQLQFC-VCCGFCGKVLVPCDMSHLANKVLEELVIGHEVTVLTHSKP 67
DB 1 VELLQLQFCVSSGSCGKVLVWPTESHMINIKILNELAQGHVETVLYSSAS 60
QY 68 KPSALKFEVHMPDORTENEIFVDLALNVLP---GLST---MOSVITKLNDFVEIRGT 120
DB 61 KSSAIKLEEV--YPTSLTK--NYLEDLSLKLIDRWIYGVSKMTFMSYFQLOELCAVYDY 116
QY 121 LKMMCESFIYNOTLKKKLOETNYDMLIDPVIIPCGDIAELLAVPFLTLRTISVGNM 180
DB 117 SNKLCKDAVYVNNKMLKMLKLOESRFDVIGDVPVPCGELLSELINLPFYVRSRISQYTIK 176
QY 181 SCGLPAPLSTVYPVMTGTDRTFLERVKNSMLSVLFHFWIQDYHFWEEFYSKALGR 240
DB 177 NGGGLFPSPSYVVALSELSDRMTFVERKNNLMILYDFWFQPFKEKSMOFSYDVLGR 236
QY 241 PTTLLETGVKAEIWLIRTYMDEFPPOPYPNFEFVGLHCKPAKALPKRMENFVSSG 300
DB 237 PTTLLETGVKAEIWLIRTYMDEFPPOPYPNFEFVGLHCKPAKALPKRMENFVSSG 296
QY 301 GIVFSLGSFQNTVEKANIATASALAOIPQKVLWRKYGKPKSTLGANTRLDWTIPON 360
DB 297 GIVFSLGSFQNTVEKANIATASALAOIPQKVLWRKYGKPKSTLGANTRLDWTIPON 356
QY 361 LGHPKRAFTTHGNGNGIYEAITYHGVPMVGPPIFGQDLNDINAMKGAAGAEINFTKMTS 420
DB 357 LGHPKRAFTTHGNGNGIYEAITYHGVPMVGPPIFGQDLNDINAMKGAAGAEINFTKMTS 416
QY 421 EDLLRALRTVITDSSYKENAMRLSRHHDDQVVKPLDRAVFWIEFVNRHKGAKHLRSAAHD 480
DB 417 EDLLRALRTVITDSSYKENAMRLSRHHDDQVVKPLDRAVFWIEFVNRHKGAKHLRSAAHD 476
QY 481 LTMFOHYSIDVIGFLTCVATAIFLTKCFLESCQKFNKTRKIEKRE 527
DB 477 LTMFOHYSIDVIGFLTCVATAIFLTKCFLESCQKFNKTRKIEKRE 523

RESULT 6